УДК 519.7:007.52; 519.711.3

И.Д. Вечирская, Г.Г. Четвериков
Харьковский национальный технический университет, Украина
Украина, 61166, г. Харьков, пр. Ленина, 14

Линейные логические преобразования в задачах семантической разметки текста

I.D. Vechirska, G.G. Chetverikov
Kharkiv National University of Radioelectronics, Ukraine
Ukraine, 61166, c. Kharkiv, Lenina av., 14.

Linear Logical Transformations is in the Semantic Marking Problem of Text

І.Д. Вечірська, Г.Г. Четвериков
Харківський національний технічний університет, Україна
Україна, 61166, м. Харків, пр. Леніна, 14

Лінійні логічні перетворення в задачах семантичної розмітки тексту

В статье исследован процесс построения цепей лексических единиц. Приведено его формальное описание с помощью теории линейных логических преобразований. Доказано утверждение об общем виде линейных логических преобразований. Проведен анализ вычисления линейных логических преобразований для n переменных. Исследованы правила вычисления в зависимости от общего вида области определения. Приведены примеры вычисления по каждому правилу.
Ключевые слова: семантическая разметка, цепочка лексических единиц, предикат, линейное логическое преобразование.

In the article the process of building chains of lexical units has been investigated. Its formal description have been presented using the theory of linear logical transformation. The theorem on general form of linear logical transformation is proved. Computation of linear logical transformation is analyzed depending on n variables. Computations rules of definitional domain are investigated depending on general form. Computation on every rule is given an example.
Key words: semantic marking, chain of lexical units, predicate, linear logical transformation.

У статті досліджено процес побудови ланцюгів лексичних одиниць. Наведено його формальний опис за допомогою теорії лінійних логічних перетворень. Доведено твердження про загальний вигляд лінійних логічних перетворень. Проведено аналіз обчислення лінійних логічних перетворень для n змінних. Досліджено правила обчислення в залежності від загального вигляду області означення. Наведено приклади обчислення по кожному правилу.
Ключові слова: семантична розмітка, ланцюг лексичних одиниць, предикат, лінійне логічне перетворення.

Введение

Разнообразие задач автоматической обработки естественно-языковых структур порождает многие актуальные на сегодняшний день вопросы. Как правило, это связано с неспособностью объективно описать субъективные состояния человеческого мезга [1-4].

Обусловлено это, с одной стороны, недостаточно глубокими исследованиями по описанию процессов обработки естественно-языковой информации человеческим мозгом, а с другой – недостаточными формальными средствами для описания естественно-языковых отношений на морфологическом, синтаксическом и семантическом уровнях. Для целей лексикографии необходимо намного более тонкое описание лексических единиц, к определению которых необходимо привлекать лингвистические понятия, связанные со строением слов, приписыванием им значений определенных грамматических категорий, происхождением, понятиями значения и смысла, функционированием в контексте и т.п. [5].

Важной задачей лексикографии является задача разметки словарей. Морфологическая разметка словаря содержит и некоторую семантическую информацию (например, признак «одушевленность» у существительных). Внедрение же семантической разметки значительно расширяет возможности лингвистов при использовании корпуса. Семантическая разметка дает возможность пользователю составлять содержательные запросы при поиске примеров употребления слова в определенном значении, а разработчикам корпуса – создавать семантические фильтры для автоматического снятия неоднозначностей [6]. На пути ее осуществления одно из основных заданий – это построение классификаторов семантических структур. Например, в [4] обозначено четыре типа отношений, каждое из которых разбивается на подклассы: отношение семонимии (синонимы, антонимы, паронимы, омонимы); отношение словообразования (слова с одинаковым корнем); тезаурусные отношения (род-вид, часть-целое, комплекс-элемент, причина-следствие); отношение ассоциаций и аналогий (ассоциаторы и аналогемы).

Достаточно удачная семантическая классификация была произведена с помощью программы ПроСеКа [7], хотя, к сожалению, протестирована на достаточно узкой предметной области (сказки). В ней реализованы для пользователя возможности задавать, редактировать и анализировать семантические отношения между лексическими единицами в виде цепочек, элементы которых связаны отношением «толкуется через», сохранять эти данные в форме, ориентированной на компьютер. Другая, разработанная в этом направлении, программа «Построение Гиперцепочек» является автоматизированной, т.е. работа по построению может выполняться пользователем, непосредственно самой программой, или же предлагается комбинированный поиск. Кроме этого, программа позволяет строить цепочки и по отношению синонимии [8]. При построении программных средств неоднократно возникали вопросы, получение ответов на которые предполагало осуществление соответсвующей заданному слову формальной постановки задачи, нахождения метода поиска, основанного на разработанных правилах их определения, который скорректировал бы критерий окончания построения цепочек лексических единиц. Поэтому важно разработать математический инструментарий для построения цепочек лексических единиц, которые в свою очередь использовались бы для задач семанической классификации.

Как правило, задачи, связанные с семантикой, имеют в большинстве своем достаточно обобщенный характер. Поэтому для построения соответствующих методов и моделей целесообразно использовать понятия и принципы достаточно высокого уровня абстракции.

Дальнейшие иседования будут базироваться на понятиях и принципах отдельных разделов математической логики, а именно на аппарате алгебры конечных предикатов.

Таким образом, **целью данной работы** является развитие средств моделирования естественно-языковых объектов для классификации семантических структур путем формального описания построения цепочек лексических единиц, а именно исследование внутренней структуры *n*-местных линейных логических преобразований и правил их вычисления.

Постановка задачи исследования

Вербальная постановка задачи была фактически приведена при анализе программы «Построение Гиперцепочек». Таким образом был описан процесс нахождения цепочек лексических единиц. На рис. 1 графически представлена схема построения цепочек лексических единиц для любого естественного языка, где x – исходное слово, x_{11} – первое характерное слово из толкования (синонимов) слова x, x_{1m} – последнее характерное слово из толкования (синонимов) слова x, x_{21}^{11} – первое характерное слово из толкования (синонимов) слова x_{11}, x_{2n}^{11} – последнее характерное слово из толкования (синонимов) слова x_{11}, x_{21}^{1m} – первое характерное слово из толкования (синонимов) слова x_{1m}, x_{2q}^{1m} – последнее характерное слово из толкования (синонимов) слова x_{1m} и так далее. Под «характерным» понимаем слово из правой части словарной статьи толкового словаря, которое непосредственно характеризирует исходную лексическую единицу, имеет то же концептуальное значение (включает такие же семантические компоненты).

Таким образом, можно записать, что x_{ij} – это j-е характерное слово на i-м уровне цепочки лексических единиц (номер лексической единицы, к которой относится толкование, определяет верхний индекс слова), где j – порядковый номер слов толкования (нахождения синонимов) лексической единицы.

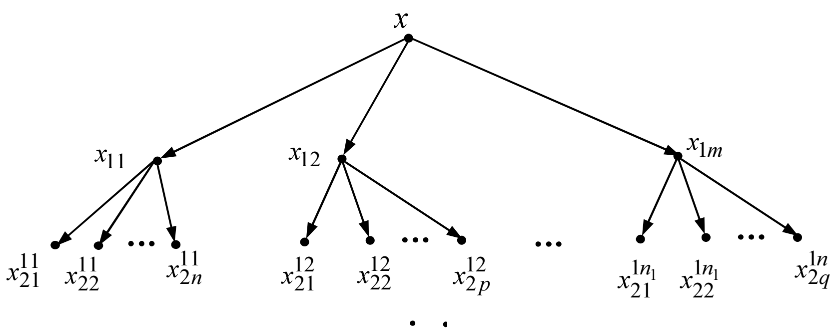


Рисунок 1 – Схема построения цепочки лексических единиц

Ранее в работе [9] было приведено формальное описание построения цепочек лексических единиц с помощью теории линейных логических преобразований. Применение метода нахождения *n*-о линейного логического преобразования и правил построения схем синтаксического подчинения позволило формализовать процесс выделения характерных слов толкования для автоматического построения цепочек лексических единиц.

В качестве лексической единицы было выбрано заголовочное слово левой части словарной статьи толкового словаря.

Однако поскольку семантическая разметка направлена на усиление возможностей поиска по лингвистическим параметрам, то представляет интерес, когда поиск в текстах ведется не только по отдельным словам, но и по их сочетаниям, т.е. конструкциям. Таким образом, исследование построения цепочек лексических единиц по словосочетанию является перспективной задачей. Поскольку сам процесс построения формально был представлен как нахождение *n*-о линейного логического преобразования, то при построении цепочки по словосочетанию, состоящему, например, из трех слов, необходимо исследовать внутреннюю структуру трехместных линейных логических преобразований для определения правил их вычисления.

Вычисление линейных логических преобразований

В статье [10] приведено и доказано утверждение об общем виде линейных логических преобразований на случай трех переменных. Рассмотрим обобщение этого утверждения на случай k переменных.

Обобщение утверждения об общем виде линейного логического преобразования на случай k переменных. Для того, чтобы функция

$$F : L_{A_{n_1} \times A_{n_2} \times \ldots \times A_{n_k}} \to P_{l_{A_l}},$$

где $L_{A_{n_1} \times A_{n_2} \times \ldots \times A_{n_k}} = \left(P_{n_1} \times P_{n_2} \times \ldots \times P_{n_k}\right)_{A_{n_1} \times A_{n_2} \times \ldots \times A_{n_k}}$ была линейным логическим преобразованием, необходимо и достаточно, чтобы она имела вид

$$[F(L)](x_l) = \bigvee_{x_{n_1}, x_{n_2}, \ldots, x_{n_k} \in S} \left(K\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}, x_l\right) P_{n_1}\left(x_{n_1}\right) P_{n_2}\left(x_{n_2}\right) \ldots P_{n_k}\left(x_{n_k}\right)\right) \qquad (1)$$

для любого $x_l \in A_l$, где $K\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}, x_l\right)$ задан на $A_{n_1} \times A_{n_2} \times \ldots \times A_{n_k} \times A_l$, $S = A_{n_1} \times A_{n_2} \times \ldots \times A_{n_k}$.

Доказательство. Достаточность. Пусть условие (1) выполнено. Тогда для любых $L, L_1, L_2 \in L_{A_{n_1} \times A_{n_2} \times \ldots \times A_{n_k}}$ и любого $\alpha \in \{0,1\}$ имеем

$$\begin{aligned}
&[\mathcal{F}(L_1 \vee L_2)](x_l) = \left| L\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}\right) = P_{n_1}\left(x_{n_1}\right) \wedge P_{n_2}\left(x_{n_2}\right) \wedge \ldots \wedge P_{n_k}\left(x_{n_k}\right) \right| = \\
&= \bigvee_{x_{n_1}, x_{n_2}, \ldots, x_{n_k} \in S} \left(K\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}, x_l\right)\left(L_1\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}\right) \vee L_2\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}\right)\right)\right) = \\
&= \bigvee_{x_{n_1}, x_{n_2}, \ldots, x_{n_k} \in S} \left(K\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}, x_l\right) L_1\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}\right)\right) \vee \\
&\vee \bigvee_{x_{n_1}, x_{n_2}, \ldots, x_{n_k} \in S} \left(K\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}, x_l\right) L_2\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}\right)\right) = [\mathcal{F}(L_1)](x_l) \vee [\mathcal{F}(L_2)](x_l)
\end{aligned}$$

для всех $x_l \in A_l$. Аддитивность доказана.

Теперь докажем однородность.

$$\begin{aligned}
&[F(\alpha L)](x_l) = \bigvee_{x_{n_1}, x_{n_2}, \ldots, x_{n_k} \in S} \left(K\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}, x_l\right) \alpha L\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}\right)\right) = \\
&= \alpha \bigvee_{x_{n_1}, x_{n_2}, \ldots, x_{n_k} \in S} \left(K\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}, x_l\right) L\left(x_{n_1}, x_{n_2}, \ldots, x_{n_k}\right)\right) = \alpha [F(L)](x_l).
\end{aligned}$$

Необходимость. Пусть для всех $L, L_1, L_2 \in L_{A_{n_1} \times A_{n_2} \times \ldots \times A_{n_k}}$ и $\alpha \in \{0,1\}$ выполнено $\mathcal{F}(L_1 \vee l_2) = \mathcal{F}(L_1) \vee \mathcal{F}(L_2)$, $\mathcal{F}(\alpha L) = \alpha \mathcal{F}(L)$. Для любого $L \in L_{A_{n_1} \times A_{n_2} \times \ldots \times A_{n_k}}$ имеем

$$L(x_{n_1}, x_{n_2}, \ldots, x_{n_k}) = P_{n_1}(x_{n_1}) P_{n_2}(x_{n_2}) \ldots P_{n_k}(x_{n_k}) = \left(\bigvee_{\alpha_{n_1} \in A_{n_1}} \left(P_{n_1}(\alpha_{n_1}) x_{n_1}^{\alpha_{n_1}} \right) \right) \wedge$$

$$\wedge \left(\bigvee_{\alpha_{n_2} \in A_{n_2}} \left(P_{n_2}(\alpha_{n_2}) x_{n_2}^{\alpha_{n_2}} \right) \right) \wedge \ldots \wedge \left(\bigvee_{\alpha_{n_k} \in A_{n_k}} \left(P_{n_k}(\alpha_{n_k}) x_{n_k}^{\alpha_{n_k}} \right) \right) =$$

$$= \left(\bigvee_{\alpha_{n_1} \in A_{n_1}} \left(P_{n_1}(\alpha_{n_1}) \varepsilon_{\alpha_{n_1}}(x_{n_1}) \right) \right) \wedge \left(\bigvee_{\alpha_{n_2} \in A_{n_2}} \left(P_{n_2}(\alpha_{n_2}) \varepsilon_{\alpha_{n_2}}(x_{n_2}) \right) \right) \wedge \ldots \wedge$$

$$\wedge \left(\bigvee_{\alpha_{n_k} \in A_{n_k}} \left(P_{n_k}(\alpha_{n_k}) \varepsilon_{\alpha_{n_k}}(x_{n_k}) \right) \right)$$

при всех $x_{n_1} \in A_{n_1}, x_{n_2} \in A_{n_2}, \ldots, x_{n_k} \in A_{n_k}$, где $\varepsilon_{\alpha_{n_1}}(x_{n_1}) = x_{n_1}^{\alpha_{n_1}}$,

$\varepsilon_{\alpha_{n_2}}(x_{n_2}) = x_{n_2}^{\alpha_{n_2}}, \ldots, \varepsilon_{\alpha_{n_k}}(x_{n_k}) = x_{n_k}^{\alpha_{n_k}}$. Тогда

$$[\mathcal{F}(L)](x_l) = \left[F\left(\left(\bigvee_{\alpha_{n_1} \in A_{n_1}} \left(P_{n_1}\left(\alpha_{n_1} \varepsilon_{\alpha_{n_1}}(x_{n_1})\right)\right)\right) \left(\bigvee_{\alpha_{n_2} \in A_{n_2}} \left(P_{n_2}\left(\alpha_{n_2} \varepsilon_{\alpha_{n_2}}(x_{n_2})\right)\right)\right) \wedge \ldots . \wedge \right.\right.$$

$$\left.\left. \wedge \left(\bigvee_{\alpha_{n_k} \in A_{n_k}} \left(P_{n_k}\left(\alpha_{n_k} \varepsilon_{\alpha_{n_k}}(x_{n_k})\right)\right)\right) \right) \right](x_l) =$$

$$= \bigvee_{\alpha_{n_1} \in A_{n_1}} \bigvee_{\alpha_{n_2} \in A_{n_2}} \ldots \bigvee_{\alpha_{n_k} \in A_{n_k}} \left(P_{n_1}(\alpha_{n_1}) P_{n_2}(\alpha_{n_2}) \ldots P_{n_k}(\alpha_{n_k}) \right) \left[\mathcal{F}\left(\varepsilon_{\alpha_{n_1}, \alpha_{n_2}, \ldots, \alpha_{n_k}} \right) \right](x_l) =$$

$$= \bigvee_{\alpha_{n_1}, \alpha_{n_2}, \ldots, \alpha_{n_k} \in A_{n_1} \times A_{n_2} \times \ldots \times A_{n_k}} \left(P_{n_1}(\alpha_{n_1}) P_{n_2}(\alpha_{n_2}) \ldots P_{n_k}(\alpha_{n_k}) K(\alpha_{n_1}, \alpha_{n_2}, \ldots, \alpha_{n_k}, x_l) \right) =$$

$$= \bigvee_{x_{n_1}, x_{n_2}, \ldots, x_{n_k} \in A_{n_1} \times A_{n_2} \times \ldots \times A_{n_k}} \left(P_{n_1}(x_{n_1}) P_{n_2}(x_{n_2}) \ldots P_{n_k}(x_{n_k}) K(x_{n_1}, x_{n_2}, \ldots, x_{n_k}, x_l) \right),$$

где положено $K(x_{n_1}, x_{n_2}, \ldots, x_{n_k}, x_l) = \left[\mathcal{F}\left(\varepsilon_{x_{n_1}} \varepsilon_{x_{n_2}} \ldots \varepsilon_{x_{n_k}} \right) \right](x_l)$ для всех $x_{n_1}, x_{n_2}, \ldots, x_{n_k} \in A_{n_1} \times A_{n_2} \times \ldots \times A_{n_k}$, $x_l \in A_l$. Утверждение доказано.

Таким образом, мы рассмотрели преобразования из (x_n, x_k) в x_l и из $(x_{n_1}, x_{n_2}, \ldots, x_{n_k})$ в x_l.

Далее рассмотрим примеры вычисления линейных логических преобразований при задании ядра преобразования различными способами.

В статье [10] представлены следующие примеры вычисления для случая трех переменных.

Возникает вопрос – как не только записать, но и вычислить такие линейные логические преобразования?

Запишем преобразование из (x_n, x_k) в x_l следующим образом: чтобы записать эту формулу в виде, удобном для вычислений, достаточно лишь выполнить опера-

цию переброски кванторов через предикат, не зависящий от переменной, стоящей под знаком квантора:

$$R(x_3) = \exists x_2 \in U\big(Q(x_2)\big(\exists x_1 \in UK(x_1, x_2, x_3)P(x_1)\big)\big).$$

Рассмотрим пример вычисления линейного логического преобразования (рис. 2) по указанной выше формуле. Пусть $K(x_1, x_2, x_3) = x_1^a x_2^b x_3^c \vee x_1^a x_2^a x_3^a$; $P(x_1) = x_1^a \vee x_1^b$; $Q(x_2) = x_2^a \vee x_2^b$; $U = \{a, b, c\}$.

Тогда $R(x_3) = x_3^a \vee x_3^c$.

Таким образом, правило вычисления линейного логического преобразования из (x_n, x_k) в x_l по аналогии с формулой (аналог повторному интегралу) будет иметь вид

$$[F(L)](x_l) = \exists x_k \in A_k\big(P_k(x_k)\big(\exists x_n \in A_n K(x_n, x_k, x_l)P_n(x_n)\big)\big),$$

где $A_k = A_n = U$.

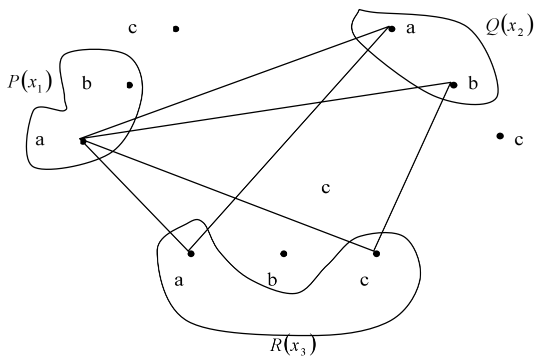


Рисунок 2 – Графическое представление
линейного логического преобразования $R(x_3)$

Далее представим правило вычисления линейного логического преобразования из $\left(x_{n_1}, x_{n_2}, ..., x_{n_k}\right)$ в x_l следующей формулой:

$$[F(L)](x_l) = \exists x_{n_k} \in A_{n_k}\left(P_{n_k}\left(x_{n_k}\right)\left(\exists x_{n_{k-1}} \in A_{n_{k-1}}\left(P_{n_{k-1}}\left(x_{n_{k-1}}\right)\right.\right.\right. ..$$

$$\left.\left.\left.\left(\exists x_{n_1} \in A_{n_1} K\left(x_{n_1}, x_{n_2}, ..., x_{n_{k-1}}, x_{n_k}\right)P_{n_1}\left(x_{n_1}\right)\right)\right)\right)\right),$$ где $A_{n_1} = A_{n_2} = ... = A_{n_{k-1}} = A_{n_k} = U$.

Мы рассмотрели простой случай, где все переменные преобразования определены на универсуме. Оказывается, при другом задании области определения, изменится и само правило вычисления. Рассмотрим далее пример (рис. 3), демонстрирующий этот факт.

$$R'(x_3) = \underset{x_1, x_2 \in F}{V} K(x_1, x_2, x_3)P(x_1)Q(x_2) = \exists x_1 \in U \exists x_2 \in U\big(F(x_1, x_2) \wedge$$

$$\wedge \big(K(x_1, x_2, x_3)P(x_1)Q(x_2)\big)\big) = \exists x_2 \in U\big(Q(x_2)\big(\exists x_2 \in UF(x_1, x_2)K(x_1, x_2, x_3)P(x_1)\big)\big).$$

Зададимся областью $F(x_1, x_2) = x_1^a x_2^a \vee x_1^b x_2^c$.

Тогда

$$R'(x_3) = \exists x_2 \in U\big(Q(x_2)\big(\exists x_2 \in F(x_1, x_2)K(x_1, x_2, x_3)P(x_1)\big)\big) = \exists x_2 \in \{a, b, c\}$$

$$\left(x_2^a \vee x_2^b\right)\left(\exists x_1 \in \{a, b, c\}\left(x_1^a x_2^b x_3^c \vee x_1^a x_2^a x_3^a\right) \wedge \left(x_1^b x_2^c \vee x_1^a x_2^a\right) \wedge \left(x_1^a \vee x_1^b\right)\right) =$$

$$= \exists x_2 \in \{a, b, c\}\left(x_2^a \vee x_2^b\right)\left(x_2^b x_3^c \vee x_2^a x_3^a\right)\left(x_2^c \vee x_2^a\right) = \exists x_2 \in \{a, b, c\}\left(x_2^a \vee x_2^b\right)x_2^a x_3^a = x_3^a.$$

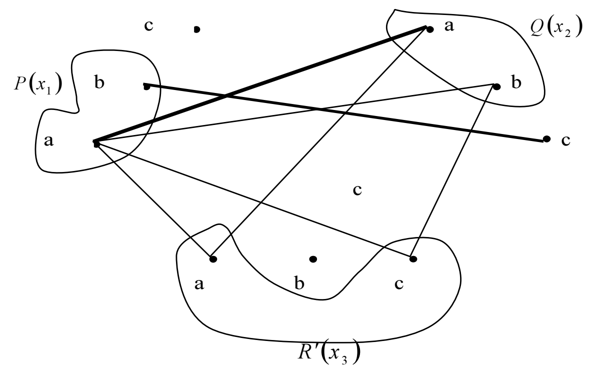


Рисунок 3 – Графическое представление линейного логического преобразования $R'(x_3)$

Таким образом, линейное логическое преобразование из (x_n, x_k) в x_l по области определения, заданной функцией $F(x_n, x_k)$, можно вычислить по следующему правилу:

$$[F'(L)](x_l) = \bigvee_{x_n, x_k \in F(x_n, x_k)} K(x_n, x_k, x_l) P(x_n) Q(x_k) = \exists x_n \in U \exists x_k \in U (F(x_n, x_k) \wedge$$
$$\wedge (K(x_n, x_k, x_l) P_n(x_n) P_k(x_k))) = \exists x_k \in U (P_k(x_k) (\exists x_n \in U F(x_n, x_k)) \wedge$$
$$\wedge K(x_n, x_k, x_l) P_n(x_n))).$$

Далее представим правило вычисления линейного логического преобразования из $(x_{n_1}, x_{n_2}, ..., x_{n_k})$ в x_l следующей формулой:

$$[F'(L)](x_l) = \bigvee_{x_{n_1}, x_{n_2}, ..., x_{n_{k-1}}, x_{n_k} \in F(x_{n_1}, x_{n_2}, ..., x_{n_{k-1}}, x_{n_k})} K(x_{n_1}, x_{n_2}, ..., x_{n_{k-1}}, x_{n_k}, x_l) \wedge$$
$$\wedge P_{n_1}(x_{n_1}) P_{n_2}(x_{n_2}) ... P_{n_{k-1}}(x_{n_{k-1}}) P_{n_k}(x_{n_k}) = \exists x_{n_1} \in U \exists x_{n_2} \in U ... \exists x_{n_{k-1}} \in U \exists x_{n_k} \in U$$
$$(F(x_{n_1}, x_{n_2}, ..., x_{n_{k-1}}, x_{n_k})(K(x_{n_1}, x_{n_2}, ..., x_{n_{k-1}}, x_{n_k}, x_l) P_{n_1}(x_{n_1}) P_{n_2}(x_{n_2}) \wedge ... \wedge$$
$$\wedge P_{n_{k-1}}(x_{n_{k-1}}) P_{n_k}(x_{n_k}))) = \exists x_{n_k} \in U (P_{n_k}(x_{n_k})(\exists x_{n_{k-1}} \in U (P_{n_{k-1}}(x_{n_{k-1}}) \wedge ... \wedge$$
$$\wedge (\exists x_{n_1} \in U (F(x_{n_1}, x_{n_2}, ..., x_{n_{k-1}}, x_{n_k})) K(x_{n_1}, x_{n_2}, ..., x_{n_{k-1}}, x_{n_k}, x_l) P_{n_1}(x_{n_1}))))).$$

Интерес представляет случай, когда $x_1 \in F_1(x_1, x_2)$ и $x_2 \in F_2(x_1, x_2)$. Пусть

$$F_1(x_1, x_2) = F(x_1, x_2) = x_1^a x_2^a \vee x_1^b x_2^c, \; F_2(x_1, x_2) = x_1^a x_2^a \vee x_1^b x_2^b.$$

В графическом виде представим соответствующее линейное логическое преобразование на рис. 4.

Формула для вычисления будет иметь следующий вид:

$$R''(x_3) = \bigvee_{\substack{x_1 \in F_1(x_1, x_2) \\ x_2 \in F_2(x_1, x_2)}} K(x_1, x_2, x_3) P(x_1) Q(x_2) =$$
$$= (\exists x_2 \in U F_2(x_1, x_2) \vee \exists x_1 \in U F_1(x_1, x_2))(K(x_1, x_2, x_3) P(x_1) Q(x_2)) =$$
$$= \exists x_2 \in U (Q(x_2)(\exists x_1 \in U (K(x_1, x_2, x_3)(F_1(x_1, x_2) \vee F_2(x_1, x_2)) P(x_1)))).$$

Тогда

$$R''(x_3) = \exists x_2 \in \{a, b, c\} (x_2^a \vee x_2^b)(\exists x_1 \in \{a, b, c\} ((x_1^a x_2^b x_3^c \vee x_1^a x_2^a x_3^a) \wedge$$
$$\wedge ((x_1^a x_2^a \vee x_1^b x_2^c) \vee (x_1^a x_2^a \vee x_1^b x_2^c)) \wedge (x_1^a \vee x_1^b))) = \exists x_2 \in \{a, b, c\} (x_2^a x_3^a \vee x_2^b x_3^c) = x_3^a \vee x_3^c.$$

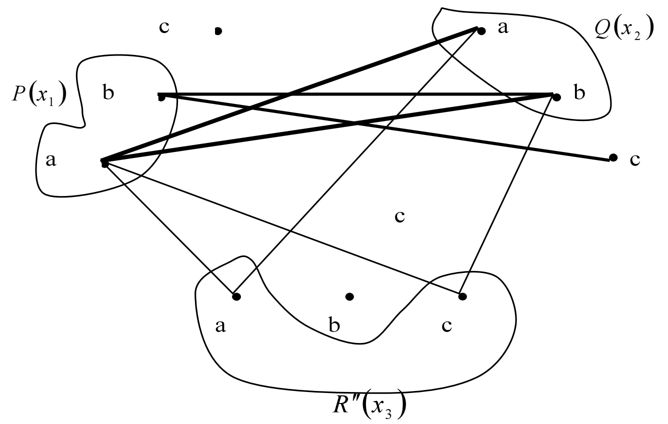


Рисунок 4 – Графическое представление линейного логического преобразования $R''(x_3)$

Таким образом, линейное логическое преобразование из (x_n, x_k) в x_l по области определения, заданной функцией $F(x_n, x_k)$, можно вычислить по следующему правилу:

$$[F''(L)](x_l) = \bigvee_{\substack{x_n \in F_n(x_n, x_k) \\ x_k \in F_k(x_n, x_k)}} K(x_n, x_k, x_l) P_n(x_n) P_k(x_k) =$$

$$= (\exists x_k \in U F_k(x_n, x_k) \vee \exists x_n \in U F_n(x_n, x_k))(K(x_n, x_k, x_l) P_n(x_n) P_k(x_k)) =$$

$$= \exists x_k \in U(P_k(x_k)(\exists x_n \in U(K(x_n, x_k, x_l)(F_n(x_n, x_k) \vee F_k(x_n, x_k)) P_n(x_n)))).$$

Далее представим правило вычисления линейного логического преобразования из $(x_{n_1}, x_{n_2}, ..., x_{n_k})$ в x_l следующей формулой:

$$[F''(L)](x_l) = \bigvee_{\substack{x_{n_1} \in F_{n_1}(x_{n_1}, x_{n_2}, ... x_{n_{k-1}}, x_{n_k}) \\ x_{n_2} \in F_{n_2}(x_{n_1}, x_{n_2}, ... x_{n_{k-1}}, x_{n_k}) \\ \dots\dots\dots\dots \\ x_{n_{k-1}} \in F_{n_{k-1}}(x_{n_1}, x_{n_2}, ... x_{n_{k-1}}, x_{n_k}) \\ x_{n_k} \in F_{n_k}(x_{n_1}, x_{n_2}, ... x_{n_{k-1}}, x_{n_k})}} K(x_{n_1}, x_{n_2}, ..., x_{n_{k-1}}, x_{n_k}) \wedge$$

$$\wedge P_{n_1}(x_{n_1}) P_{n_2}(x_{n_2}) ... P_{n_{k-1}}(x_{n_{k-1}}) P_{n_k}(x_{n_k}) = (\exists x_{n_k} \in U F_{n_k}(x_{n_1}, x_{n_2}, ... x_{n_{k-1}}, x_{n_k}) \vee$$

$$\vee \exists x_{n_{k-1}} \in U F_{n_{k-1}}(x_{n_1}, x_{n_2}, ... x_{n_{k-1}}, x_{n_k}))(K(x_{n_1}, x_{n_2}, ..., x_{n_{k-1}}, x_{n_k}) \wedge$$

$$\wedge P_{n_1}(x_{n_1}) P_{n_2}(x_{n_2}) ... P_{n_{k-1}}(x_{n_{k-1}}) P_{n_k}(x_{n_k}) = \exists x_{n_k} \in U(P_{n_k}(x_{n_k}) \wedge$$

$$\wedge (\exists x_{n_{k-1}} \in U(P_{n_{k-1}}(x_{n_{k-1}}) ... (\exists x_{n_1} \in U(K(x_{n_1}, x_{n_2}, ..., x_{n_{k-1}}, x_{n_k}) \wedge$$

$$\wedge (F_{n_1}(x_{n_1}, x_{n_2}, ... x_{n_{k-1}}, x_{n_k}) \vee F_{n_2}(x_{n_1}, x_{n_2}, ... x_{n_{k-1}}, x_{n_k}) \vee ... \vee$$

$$\vee F_{n_{k-1}}(x_{n_1}, x_{n_2}, ... x_{n_{k-1}}, x_{n_k}) \vee F_{n_k}(x_{n_1}, x_{n_2}, ... x_{n_{k-1}}, x_{n_k})) P_{n_1}(x_{n_1})))).$$

Выводы

Таким образом в статье был проведен анализ процесса построения цепочек лексических единиц.

Для его формализации выбран аппарат алгебры конечных предикатов, а именно линейные логические преобразования. С помощью линейных логических преобразований было осуществлено математическое описание самого процесса построения.

Был проведен детальный анализ действий над линейными логическими преобразованиями. Исследовано на разных областях (когда области определения зависят от одной и/или двух переменных) вычисление линейных логических преобразований. Полученные правила вычисления линейных логических преобразований проиллюстрированы примерами для трех переменных.

Перспективы дальнейших исследований. Как известно, основной структурной единицей словаря служит словарная статья. Каждая словарная статья содержит регистровую единицу – заголовочное слово, которое является своеобразным идентификатором статьи. Разным заголовочным словам соответствуют разные словарные статьи. В свою очередь статьи с одним и более регистровым словом объединяют по тождественному набору лексических значений в регистровый ряд. Регистровыми единицами могут служить лексические единицы, образующие компоненты составных слов, слова, употребляемые только как компоненты установившихся словосочетаний [11]. Для исследования внутренней структуры регистрового ряда можно использовать правила вычисления линейных логических преобразований.

Что касается построения цепочек лексических единиц, то еще одну перспективную задачу, которую можно исследовать с помощью полученных нами правил, сформулируем в следующем виде. Левая часть полноструктурированной словарной статьи содержит такие элементы: регистровый ряд; показатели словоизменительной парадигмы; показатели синтаксических связей и функций; грамматические параметры (род для существительных, вид для глаголов, показатель части речи, что указывается для местоимений, наречий, числительных, служебных частей речи и восклицаний); стилистические и другие ремарки. Таким образом, интересной и несомненно перспективной для автоматической реализации является задача построения семантической цепочки по всей левой части словарной статьи, где будут учитываться разные задания областей определения для составляющих левой части словарной статьи на основе исследованных правил вычисления линейных логических преобразований.

Литература

1. Бондаренко М.Ф. Мозгоподобные структуры : [справочное пособие] / М.Ф. Бондаренко, Ю.П. Шабанов-Кушнаренко. – К. : Наукова думка, 2011. – Том первый : [текст]. – 460 с.
2. Морковкин В.В. О единицах лексической системы : [текст] / В.В. Морковкин // Лексика и лексикография : сб. науч. трудов / [отв. ред. Ю.Г. Коротких, А.М. Шахнарович]. – М., 1992. – С. 127-134.
3. Бондаренко М.Ф. Концепції уніфікації інформаційно-інтелектуальних технологій в системах мовлення : [текст] / М.Ф. Бондаренко, З.Д. Коноплянко, Г.Г. Четвериков // Біоніка інтелекту: наук.-техн. журнал. – 2011. – № 3 (77). – С.150-156.
4. Широков В.А. Комп'ютерна лексикографія : [текст] / Широков В.А. – Київ : Науково-виробниче підприємство «Видавництво "Наукова думка" НАН України», 2011. – 351 с.
5. Широков В.А. Елементи лексикографії : [текст] / Широков В.А. – Київ : Довіра, 2005 – 304 с.
6. Кустова Г.Н. Национальный корпус русского языка: семантические фильтры для разрешения многозначности глаголов : [текст] / Г.Н. Кустова, С.Ю. Толдова // Труды междунар. конф. : «Корпусная лингвистика – 2008». – СПб., 2008.
7. Рафаєва А.В. Програма семантичної класифікації лексики ПроСеКа : [текст] / А.В. Рафаєва // Прикладна лінгвістика та лінгвістичні технології : MegaLing-2009 : зб. наук. праць / НАН України. Укр.мовн.-інформ.фонд, Таврійськ. нац. ун-т ім. В.І. Вернадського/ [за ред. В.А. Широкова]. – К. : Довіра, 2009. – 527 с.
8. Федорова Т.Н. Автоматизация построения цепочек лексических единиц на примере украинских народных сказок : [текст] / Т.Н. Федорова [Электронный ресурс] // Материалы международной научной конференции : «Горизонты прикладной лингвистики и лингвистических технологий»

(MegaLing'2011). – 20 – 26 сентября 2011, Украина, Киев. – Режим доступа : http://megaling.ulif.com.ua/.
9. Вечірська І.Д. Математичні аспекти побудови ланцюгів лексичних одиниць : [текст] / І.Д. Вечірська, Г.Г. Четвериков // Бионика интеллекта: науч.-техн. журнал. – 2012. – № 2(79). – С. 84-88.
10. Вечирская И.Д. О вычислении линейных логических преобразований / И.Д. Вечирская, А.А. Иванилов // Вестник НТУ «ХПИ» : сб. науч. трудов. – Харьков : НТУ «ХПИ», 2005. – № 18. – С. 29-32.
11. Широков В.А. Лінгвістичні та технологічні основи тлумачної лексикографії : [текст] / В.А. Широков – Київ : Довіра, 2010 – 295 с.

Literatura

1. Bondarenko M.F. Brain-like structure: spravochnoe posobie. Tom pervyj [Text] / M.F. Bondarenko, Yu.P. Shabanov-Kushnarenko – K.: Naukova dumka, 2011. – 460 s.
2. Morkovkin V.V. About units of lexical system [Text] / V.V. Morkovkin // Leksika i leksikographiya: Sb. nauch. trudov / Otv. red. Yu.G. Korotkih, A.M. Shahnarovich – M., 1992. – S. 127-134.
3. Bondarenko M.F. Unification conception of information-intelligence technology in language systems [Text] / M.F. Bondarenko, Z.D. Konoplyanko, G.G. Chetverikov // Bionics of Intelligence: Sci. Mag. – 2011. – № 3 (77). – S. 150-156.
4. Shyrokov V.A. Computer lexicography [Text] / V.A. Shyrokov – Kyjiv: Naukova dumka, 2011. – 351 s.
5. Shyrokov V.A. Elements of lexicography / [Text] / V.A. Shyrokov – Kyjiv: Dovira , 2005 – 304s.
6. Kustova, G.N. National corps of Russian language [Text] / G.N. Kustova, S.Yu. Toldova // Proc. «Corps linguistics – 2008». – SPb, 2008.
7. Rafayeva A.V. Program of semantic classification of vocabulary ProSeKa [Text] / A.V. Rafayeva // MegaLing-2009: Zb. nauk. prats / Za red. За ред. V.A. Shyrokova – Kyjiv: Dovira, 2009. – 527 s.
8. Fedorova T.N. Automatization of building of chain of lexical units on example ukrainian stories [Text] / T.N. Fedorova // MegaLing-2011. – 20-26 sentyabrya 2011, Ukraina, Kyjiv, http://megaling.ulif.com.ua/
9. Vecvirska I.D. Mathematical aspects of building of chain of lexical units [Text] / I.D. Vechirska, G.G.Chetverikov// Bionics of Intelligence: Sci. Mag. – 2012. – № 2(79). – S. 84-88.
10. Vecvirska I.D. About computation of linear logical transformations / I.D. Vechirska, A.A. Ivanilov // Vestnik NTU «KHPI» – Kharkov: NTU «KHPI», 2005. – № 18. – S. 29-32.
11. Shyrokov V.A. Linguistic and technological basis of explanatory lexicography [Text] / V.A. Shyrokov – Kyjiv: Dovira, 2010 – 295 c.

RESUME

I.D. Vechirska, G.G. Chetverikov

Linear Logical Transformations is in the Semantic Marking Problem of Text

In given article, a perspective analysis of semantic marking problem were developed It directs to amplification of retrieval features by linguistic characteristics. The retrieval is realized not only by word, but by word-combination too. It is envisaging further development problem.

An approach to the construction of semantic chains for phrases was developed based on formalized in [9] of the process of building chains of lexical units. For studies it was selected the unit of the algebra of finite predicates, namely the linear logic of transformation. The linear logical transformation helped to carry out a mathematical description of the process of construction.

It was carried out a detailed analysis of the actions of the linear logic transformations. Also it was investigated a calculation of the linear logical transformation in different areas. The resulting calculation rules of linear logic transformations were illustrated with examples for three variables.

Статья поступила в редакцию 09.04.2013.